

DOCUMENT RESUME

ED 257 096

CS 208 929

AUTHOR Takala, Sauli; Vahapassi, Anneli
TITLE International Study of Written Composition.
PUB DATE 85
NOTE 20p.; Paper presented at the Annual Meeting of the International Writing Convention (Norwich, England, March 31-April 4, 1985). Figure 4 may not be legible due to small print.
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Elementary Secondary Education; Foreign Countries; *Grading; *Holistic Evaluation; *International Educational Exchange; Writing (Composition); *Writing Evaluation; Writing Processes; *Writing Research
IDENTIFIERS *Writing Tasks

ABSTRACT

Fifteen member countries of the International Association for the Evaluation of Educational Achievement (IEA) collaborated on a project that had four aims: (1) to contribute to the conceptualization of the domain of writing, (2) to develop an internationally appropriate set of writing tasks and a system for assessing compositions, (3) to describe recent developments and the current state of instruction in written composition, and (4) to identify factors that explain differences and patterns in the performance of written composition. A set of writing tasks was presented in each country to one or more of the following population groups: (1) the end of primary education (Population A), (2) the end of compulsory education (Population B), and (3) the preuniversity year (Population C). The recommended sample sizes were 50 classes for Populations A and C, and 100 classes for Population B. Problems relating to the construction of writing tasks, the allocation of writing tasks, and the scoring of student scripts were addressed. The most appropriate scoring system combined a holistic, overall-impression marking and an analytical marking. It was concluded that in spite of a great effort the scoring metric was too elastic to allow robust cross-national comparisons of scores on any one task or group of tasks. A sample scoring sheet, tables, and diagrams are included. (DF)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED257096

BEST COPY AVAILABLE

S 208929



INTERNATIONAL STUDY OF WRITTEN COMPOSITION

Sauli Takala & Anneli Vähäpassi
Institute for Educational Research
University of Jyväskylä

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

This document has been reproduced as
received from the person or organization
originating it.

Minor changes have been made to improve
reproduction quality.

Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Sauli Takala

Anneli Vähäpassi

INTRODUCTION

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

In this paper we will give a brief account of an on-going international study of written composition. Fifteen member countries of the IEA (The International Association for the Evaluation of Educational Achievement) from all over the world are collaborating on the project. A fairly large set of writing tasks has been presented in some countries to one population of students, in some to two and in some to all three populations, representing end of primary education (Pop. A), end of compulsory education (Pop B), and the pre-university year (Pop C), respectively (for a detailed account, see Table 1, Appendix 1). The recommended minimum sample sizes were 50 classes for Populations A and C and 100 for Pop B.

We will first place the study within its larger context. After that we proceed to describe the aims of the project, its design and scope. We will give particular emphasis to the tasks and to the scoring procedures. We will also present some results concerning the pilot stages of the project.

We wish to point out that this is not the "official" or complete presentation of the IEA Written Composition Study. We have chosen to highlight some aspects of the study and have added some personal reflections and interpretations.

Paper presented at an International Writing Convention, School of Education, University of East Anglia, Norwich, March 31-April 4, 1985.

CONTEXT OF THE STUDY

Since the late 1950's, a number of educational researchers and research institutions have been working on an empirically oriented comparative research program. A small feasibility study was carried out by the IEA in 1959-1962. This was followed by a First Mathematics Study in the mid 1960's (with 12 countries/ school systems participating) and by six studies in 1970-1971. These covered Science (19 countries), Reading Comprehension (15), Literature (10), French as a foreign language (8), English as a foreign language (10), and Civic Education (10). On-going studies comprise a second mathematics and science study and a study of classroom activities. Since 1979, work has been carried out on an international study of written composition.

A number of international and national reports have been published on the completed studies. Similarly, a great number of journal articles have been published, as a bibliography of IEA-related publications shows.

MANAGEMENT OF THE STUDY

Like all IEA studies, the Written Compositions study has a complex management structure. The International Project Council at its annual meeting makes general policy decisions. Dr. Alan C. Purves from the University of Illinois at Urbana-Champaign is the Chair of the IPC. More specific planning is the responsibility of the International Steering Committee, chaired by Anneli Vähäpassi from Finland. Members are Tom Gorman (England & Wales), Judit Kadar-Fulop (Hungary), Eva Baker (USA), Alan C. Purves (USA), Hildo Wesdorp (Netherlands, until 1985), Pai Obanya (Nigeria) and Raimo Konttinen (Finland). Sauli Takala (Finland) is the International Coordinator and Elaine Degenhart (USA) Deputy Coordinator.

From January 1981 to fall 1984, the International Coordinating Center was located at the Curriculum Laboratory (UIUC). Since the fall of 1984 the coordinating of the project is managed jointly by the Curriculum Laboratory and the Institute for Educational Research, University of Jyväskylä.

National Research Coordinators constitute the International Study Committee, which has met a few times to discuss the implementation of the study according to common plans.

The costs of international coordination have been paid by the IEA, while the national costs of implementing the study are paid by each participating country.

DESIGN AND AIMS OF THE STUDY

The IEA International Study of Written Composition seeks to accomplish the following tasks:

- (1) to contribute to the conceptualization of the domain of writing and particularly the domain of school-based written composition,
- (2) to develop an internationally appropriate set of writing tasks and a system for assessing compositions which is applicable across countries/school systems and across languages,
- (3) to describe recent developments and the current state of instruction in written composition in the participating countries/school systems, and
- (4) to identify factors which explain differences and patterns in the performance of written composition and other outcomes, with particular attention to cultural background, curriculum and teaching practices.

The independent and dependent variables of the study and their presumed relationships are illustrated in Figure 1. We will

=====
Place Figure 1 about here
=====

BEST COPY AVAILABLE

begin with the conceptualization and operationalization of the dependent variable.

DEPENDENT VARIABLE: WRITING TASKS AND SCORES

Problems related to the dependent variable

The greatest challenge in the study has been the work on the dependent variable. Here we can give only a brief account of it (for a detailed account, see Vähäpassi, 1982, 1983). Three problem areas had to be addressed:

I. Problems related to the construction of writing tasks

- (a) What is the total domain of writing, especially school-based domain of written composition? This required conceptual analysis and synthesis.
- (b) What is the appropriate sample from the total domain for the students concerned? What sub-domains should be included in the set of writing tasks?
- (c) What is the appropriate system for specifying the tasks?
- (d) How should the actual writing tasks be formulated?

II. Problems related to the allocation of writing tasks

- (e) What kind of tasks should be presented to each of the three student populations?
- (f) How should populations be linked through common tasks?
- (g) How many tasks should/can each student be asked to respond to?
- (i) If task rotation is necessary, how should it be done so as to maximize the information obtained and minimize problems that are related to rotation?

III. Problems related to rating of student scripts

- (j) What rating system should be used (e.g., holistic, primary trait, analytical)?
- (k) How can a rating system be constructed which can be applied in a comparable way in all participating countries?

It is possible to focus here on a few of these problems only.

Specification of the Domain of Writing

After an extensive review of the problem area and after having tried several systems and discarded them, the following approach (Figure 2) was adopted (for a more detailed account; see Vähäpassi, 1982, 1983).

=====
Place Figure 2 here
=====

The system relates the primary purpose and audience of writing to the level and object of cognitive processes involved in the writing task. This gives us a typology and we can test it in some respects by trying to place various communication tasks and writing products in its cells. This has been illustrated selectively in Figure 2. Without going into detail; suffice it to say that up to now, the model has served as a very useful tool in solving a number of problems. We will return to the model later on.

Sampling of tasks from the domain

In the IEA study, writing has been considered an act of communication about some topic between the writer and the reader(s) (addressee/audience). The writer has a certain purpose in writing and his ultimate goal is communicative effectiveness (effective expression). Typically the writer produces a text alone, without the immediate feedback from (cooperation/negotiation with) the reader. This latter aspect seems to be one of the most important features that distinguishes writing from face-to-face conversational interaction (Takala, 1982, 1983). The above discussion has been shown schematically in Figure 3.

=====
Place Figure 3 here
=====

The tasks of the IEA Written Composition Study can be related to the domain specification grid (Figure 2) and to the rhetorical model of writing (Figure 3) as follows:

(1) Tasks that emphasize the perspective of the writer

Task 5: write a personal story

Task 8: write a "free" composition on an ambiguous and evocative pictorial stimulus

(2) Tasks that emphasize the perspective of the topic

Task 2: summarize a text

Task 3: retell a story (in a shorter form)

Task 4a: describe a ritual mask

Task 4b: describe a process of doing something

Task 7: write a reflective essay

(3) Tasks that emphasize the perspective of the reader

Task 6: try to persuade the reader to share the writer's strong view about something

(4) Tasks that have several perspectives

Task 1a: describe a desired bicycle to an uncle who wishes to buy one as a birthday present

Task 1b: describe oneself to a penfriend whom the student is going to visit so as to make it possible for the penfriend to identify the student as he comes to meet the student

Task 1c: write a note to the principal/headmaster canceling a scheduled meeting

Task 1d: leave a message at home telling where the student has gone after school

Task 1e: write a letter applying for an advertised summer job

Note that the writer-reader social status and the topic is varied systematically.

Task 9: write a letter to a younger student who is coming to study at the same school as the writer, telling the new student how he/she should write in the new school to get good grades.

If we focus on the purpose of the tasks, we can see that Tasks 1, 2, 3, 4 and 9 have a predominantly informational purpose. The expressive purpose is dominant in Tasks 5 and 8.

Tasks 1e and 6 have a persuasive purpose. Task 7 has an explicative/interpretive purpose.

If we pay attention to the cognitive structure of the tasks, we might suggest that Tasks 1c, 1d, 2, 3, and 5 have a structure based mainly on a temporal organization (an account of events). Tasks 1a, 1b and 4a have a structure based mainly on a spatial organization (an account of the physical characteristics of objects). Tasks 1e, 6, 7, 8 and 9 have a structure based mainly on a logical organization (an account of ideas and thought structures). Other classifications are, of course, possible. For an example of actual tasks, see Appendix 2.

Scoring system

Scoring scheme

After considering and trying out various scoring systems (holistic, primary trait; analytical), it was agreed that the most appropriate one was a system which combined a holistic overall impression marking and analytical marking. Figure 4 is an attempt to summarize the conceptualization work done in developing the scoring system.

=====
Place Figure 4 here
=====

The overall construct "student writing" is divided into two sub-constructs: "writing competence" and "writing preferences". We are concerned here mainly with writing competence, which is taken to consist of "discourse-structuring competence" (or rhetorical competence) and of "text-producing competence". Rhetorical competence is assumed to require both "cognitive competence" and "social competence".

BEST COPY AVAILABLE

Cognitive competence is manifested in illocutionary effectiveness, i.e., in the ability of writers to make readers easily recognize the communicative intent. This is operationalized in our scoring scheme as rated "Quality and scope of ideas" and as rated success in the "Organization and Presentation of Content". In some tasks, writers also aim at perlocutionary effectiveness, i.e., they want to get the readers to do something. This may mean that writers wish to have readers entertained or have them change their minds, or even do something concrete after such a change of mind. Such perlocutionary effectiveness presupposes awareness of readers' norms, attitudes, etc. In our scoring scheme this is operationalized, in part, as "Style and Appropriateness of Tone".

Text-producing competence presupposes "linguistic competence" and "motor competence". These are operationalized in our scoring scheme as rated competence in using "grammatical features", "spelling and orthographic conventions", and "handwriting and neatness", all rated in accordance with national norms. For a sample scoring scheme, see Appendix 3.

The fact that raters are readers and will, in spite of their attempt to rate compositions as objectively as possible, according to the categories of the scoring scheme, have a subjective response to the scripts. For this reason, we have added a category, "Response of Rater", which allows raters to indicate such a response. This is another way of trying to get an estimate of perlocutionary effectiveness.

Scoring scales

The scoring scheme is supplemented with a scoring scale, which consists of a number of compositions that exemplify various scale point values of the scheme. There were no particular problems in having a consensus on the adopted scoring scheme. On

the other hand, the construction of the scoring scales was a difficult and laborious task.

There were several problems that had to be addressed. The most difficult one was how to establish internationally agreed scores for compositions written in several different languages. A number of alternatives were considered (e.g., translation into several languages, use of bilingual raters, having an international rating panel). After long discussions, it was agreed that the most feasible solution was as follows: (1) Pilot test the tasks: (2) National Centers select a set of compositions representing the whole performance range, and rate them according to national norms. (3) Translate these proposed "benchmark" compositions carefully into English, following a common set of guidelines. (4) Assemble juries with an international composition and have them rate the (sometimes translated) compositions after the juries have been trained according to the established training system. (5) In rating the compositions of the main study, use the original benchmarks (written in the source language) with their international scores in rater training and as a reference in final rating. (6) Check the comparability of rating in participating countries by means of a common "international calibration set", which is a set of compositions with a high degree of consensus.

Some Results

The system described in the above was applied for the first time during the third International Study Committee Meeting in Urbana in the spring of 1984.

Four international juries consisting of five National Research Coordinators were set up. Members of the Steering Committee served as jury leaders and native speakers of English as

rapporteurs,

The establishment of benchmarks and calibration sets was limited mainly to Population B. Altogether some 1,400 compositions had to be rated, discussed, and commented on. Each jury member scored about 500 compositions. Jury members knew only the target aspect that had been given, a nationally proposed initial score. The compositions were handled in the following way:

- if there was a 4 out of 5 agreement on the nationally proposed target aspect, the score was internationally validated (whether it was identical with or lower or higher than the national suggestion)
- if there was a lower agreement - or if a jury wanted another jury to have a look at some compositions for whatever reasons - such compositions were referred to rescoring
- if some compositions received highly uniform scores on all target aspects, they were singled out as candidates for the international calibration set.

The results of the Urbana scoring session have been partly analyzed and can be briefly summarized as follows:

(1) By using a common set of compositions scored by all jury members after they have finished scoring a particular task, it was found that there was full agreement in 22% of all cases, 78.2% agreement with one scale point deviance. This was lower than we had hoped, but not unusual even when raters belong to the same interpretive community. The produced benchmark scores do not constitute as good a yardstick as is desirable.

(2) There was some systematic difference in jury leniency (mean of all differences in means between all pairs of juries .326).

(3) The average standard deviations of juries varied only slightly.

(4) The perception of quality of compositions was to a large extent convergent between juries as median correlations ranging between .83 and .91 on different aspects show.

(5) Jury homogeneity was fairly good (.85 or higher, Cronbach's

alpha).

(6) There was considerable scale shrinkage: the 5-point scale shrank to a 3-point scale in about half of the cases. This is to some extent an artifact, because in some cases countries had a small pilot sample and reported that they sometimes felt obliged to propose a 5 even if they felt that that was too high.

(7) More problematic are the 10% of all cases when the international jury reversed the scale (e.g., giving 2-5-4 or 3-2-4 when the proposed order of merit was 1-3-5).

The results of the second scoring meeting held in Frascati late 1984 have not been fully analysed yet.

The implications of the results were discussed at a meeting in Hamburg in February, 1985. It was concluded that in spite of a great effort the scoring metric is too elastic to allow robust cross-national comparisons of scores on any one or group of tasks. Since there is quite a lot of interest in some international comparison, some further work is needed. There are plans to try to improve the comparability using one task and one population only. A robust international scale is created and a subsample of compositions are rescored. This one task will then be used for limited international comparisons.

For anyone working in the assessment of written composition the results on an international attempt to create a fully comparable scoring procedure are not surprising. In fact, we have been somewhat surprised that we have been able to solve as many problems as is the case. While robust comparisons of achievement across countries are doubtful, we do not consider that we have failed in our attempt. It appears that tasks 1, 2, and 3 (see Design and Aims of the Study above) will be successfully accomplished. That is more than some of us dared to expect.

INDEPENDENT VARIABLES

Figure 1 presents the major constructs that underly the large set of variables used as independent variables. Detailed information has been gathered by questionnaires, interviews and written case studies about the context of written composition teaching.

The gathered information (covering hundreds of variables) will be used mainly to provide detailed descriptions of the context and status of writing instruction in participating countries. A smaller subset of variables will probably be used in some multivariate analyses and model testing.

We are particularly interested in the portraits of writing instruction that the very specific Teacher Questionnaire data will provide.

REPORTING OF RESULTS

There will be both international and national reporting on the findings of the study. Three international reports are planned. The first volume will deal with the dependent variable. It will give a detailed account of the problems and issues involved in constructing an international set of writing tasks and in scoring student scripts using an internationally agreed-upon scoring procedure. The report is expected to be in manuscript form by the end of 1985.

The second volume will give a detailed description of the context and practices of writing instruction in the participating countries.

The third volume will present the main results of the study. This will probably be a set of parallel national portraits with some international comparisons. The extent of comparisons will

depend on the degree to which the scoring is, in fact, comparable in the participating countries.

SOME CONCLUDING REMARKS

In conclusion, we would like to make some personal observations. The IEA writing study has been the most difficult and laborious task ever undertaken by either of us. It has forced us to tackle many problems without being able to draw on any good models. The pressure of deadlines has also been very heavy.

On the other hand, the IEA study has been more rewarding than any other project we have been working on. Facing new problems has been a constant challenge and given us an opportunity to learn a lot. The chance to work on our problems with dedicated and experienced colleagues from all over the world has been invaluable in finding justified and acceptable solutions. It has also taught us to be more sensitive to cultural similarities and differences. We have come to be more aware of the variety of ways in which writing instruction can be arranged.

On the basis of the enthusiastic reception the IEA study has had in our country, we are convinced that the project will have a major impact on the teaching of compositions in Finland. Through our participation in the IEA study, we in Finland will get a first good national assessment of writing. In addition to that, we will be able to relate our situation to that in several other countries. We have established close contacts with many colleagues in other countries.

There are plans to store a representative sample of student scripts in an international student text corpus. Several signs indicate that there is growing interest to move from the assessment stage to a stage, when we can take a close look at the compositions themselves. A number of cognitive, linguistic,

rhetorical, cross-cultural etc studies have been tentatively sketched during our recent meetings.

REFERENCES

Takala, S. (1982). On the origins, communicative parameters and processes of writing. *Evaluation in Education*, 5(3), 209-230.

Takala, S. (1983). The domain of writing. In S. Takala & A. Vähäpassi, *On the specification of the domain of writing* (pp. 1-61). University of Jyväskylä, Institute for Educational Research, Report No. 333.

Vähäpassi, A. (1982). On the specification of the domain of school writing. *Evaluation in Education*, 5(3), 265-289.

Vähäpassi, A. (1983). On the specification of the domain of school-writing. In S. Takala & A. Vähäpassi, *On the specification of the domain of writing* (pp. 63-93). University of Jyväskylä, Institute for Educational Research, Report No. 333.

BEST COPY AVAILABLE

TABLE 1. Tested tasks per population in participating countries (plus some background information).

Grades, Ages, Median Class Sizes and Testing Dates for Tested Populations																				
Populations to be Tested			Median Age			Year in School/Grade			Median Class Size			Tested tasks/population								
A	B	C	A	B	C	A	B	C	A	B	C	1a	1b	1c	1d	1e	2	3	4a	4b
Australia		C			16-18			12			25					C				
Chile	B			15-18				9			35	B	B	B	B				B	B
England	B			15-18				10			27	B	B	B	B				B	B
Federal Republic of Germany	B			16-18				11			28	B	B	B	B				B	B
Finland	A	B	C	12-13	15-16	17-19	6	9	12	35	32	37	A/B	A/B	A/B	A	B/C	B	A	A/B
Hungary	B	C		13-14	17-18			8	12	35	30		B	B	B	B/C			B	B
Indonesia	A			11-15			6			30			A	A	A	A			A	A
Italy	A	B	C	11	14	17-18	5	8	13	25			A/B	A/B	A/B	A	B/C	B/C	A	A/B
Netherlands	B			15-17				10		28			B	B	B	B			B	B
New Zealand	A	B	C	12	15	17		11	13	30	28		A/B	A/B	A/B	A	B/C		A	A/B
							(67)	(610)	(612)										A/B	A/B/C
Nigeria	A	B	C	10-12	15-19	17-22	6	11	13	50	45	25	A/B	A/B	A/B	A	B/C	B	A	
Sweden	A	B	C	12-13	15-18	17-18	6	9	11-12				A/B	A/B	A/B	A	B/C		A	
Thailand			C			17-18		12		35						C	C			
United States	A	B	C	11-13	15-16	16-17	9	11	11-12	25	25	25	A/B	A/B	A/B	A	B/C	B/C	A	A/B

International Study of Achievement in Written Composition

Task 6

Think of an issue or problem about which you have a strong opinion. It could be about some aspect of your home, school, community or society as a whole.

Write an essay which would persuade somebody who does not share your opinion to change his or her mind and see things your way.

Your writing will be judged on what you have to say, how well you organize and present your point of view and how clearly and effectively you express yourself.

Note: When you have finished writing, answer the questions on page 5.

Space for planning (notes, outline, draft ...)

For official use only

COUNTRY	01-02	03-04	05	06-07	08-10	11-12	13-15	16-18	19
STUDY									
POPULATION									
STRATUM									
SCHOOL									
CLASS									
STUDENT ID									
INSTRUMENT									
CARD									

BEST COPY AVAILABLE

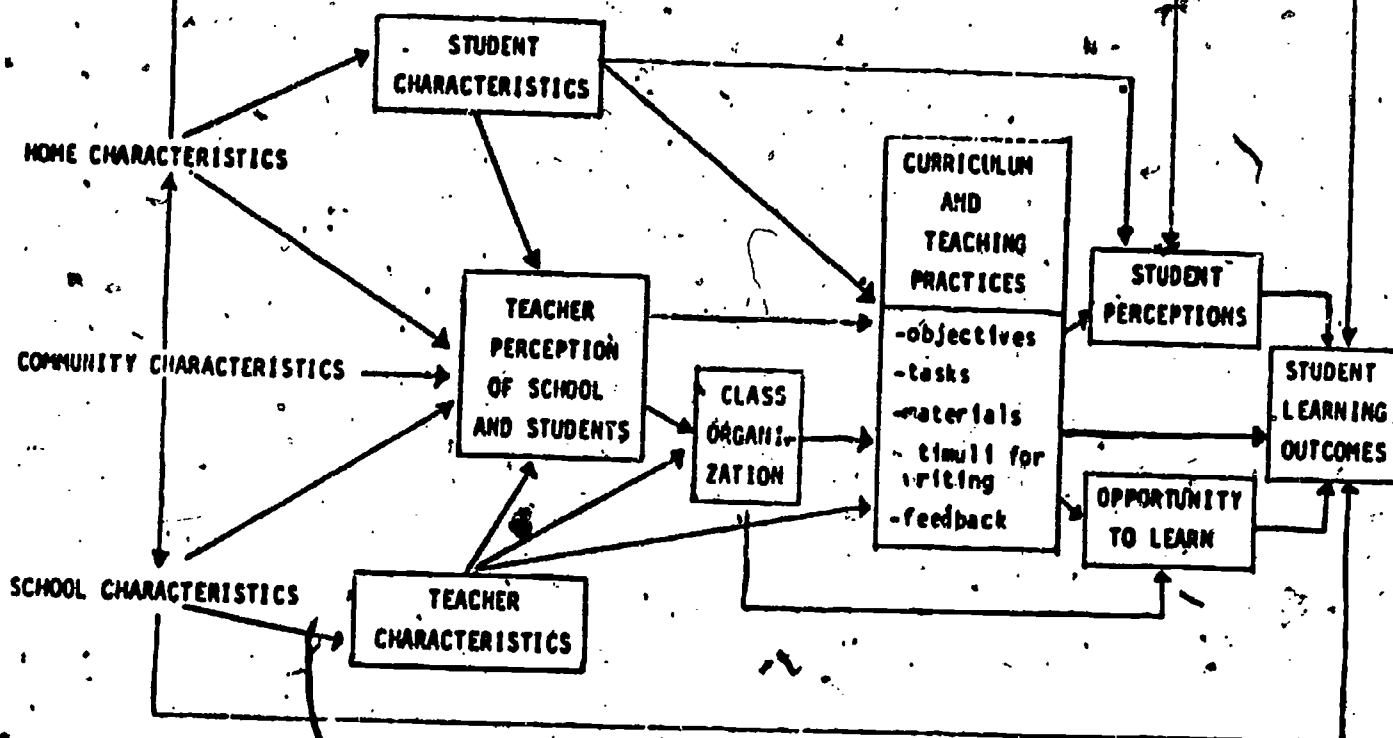


FIGURE 1. General Model of the Study (Main Constructs)

FIGURE 2. Tasks in Relation to the Domain of School Writing

Cognitive Processing		REPRODUCE	ORGANIZE/REORGANIZE	INVENT/GENERATE
Dominant Intention/Purpose	Primary Audience	Facts Ideas	Visual images, facts, Events mental states, ideas	Ideas, mental states alternative worlds
To learn (meta-lingual)	Self	Copying Taking Dictation	Retell a story (heard or read) ③ Note Resume Summary ② Outline Paraphrasing	Comments on book margins Metaphors Analogies
To convey emotions, feelings (emotive)	Self Others	Stream of consciousness	Personal story ⑤ Portrayal Personal diary Personal letter	Reflective writing --Personal essays ⑦
To inform (referential)	Others	Quote Fill in a form Message ①	Narrative report ⑨ News Instruction ⑨ Telegram Announcement Circular Message ① Directions ① Description ④ Technical description ③ Biography Science report/experiment	Expository writing --Definition --Academic essay/article ⑦ --Book review --Commentary
To convince/persuade (conative)	Others	Citation from authority/expert	Letter of application Advertisement Letter of advice ⑨ Statement of personal views, opinions ⑧	Argumentative/persuasive writing ⑥ --Editorial --Critical essay/article ⑦
To entertain, delight, please (poetic)	Others	Quote poetry and prose	Given an ending, create a story Create an ending Retell a story Word portrait or sketch Causerie ⑧	Entertainment writing --Parody ⑪ --Rhymes
PRIMARY MODE OF DISCOURSE		DOCUMENTATIVE DISCOURSE	CONSTATIVE DISCOURSE Narrative Explanatory Descriptive	EXPLORATORY DISCOURSE Interpretive (Expository/Argumentative) Literary

BEST COPY AVAILABLE

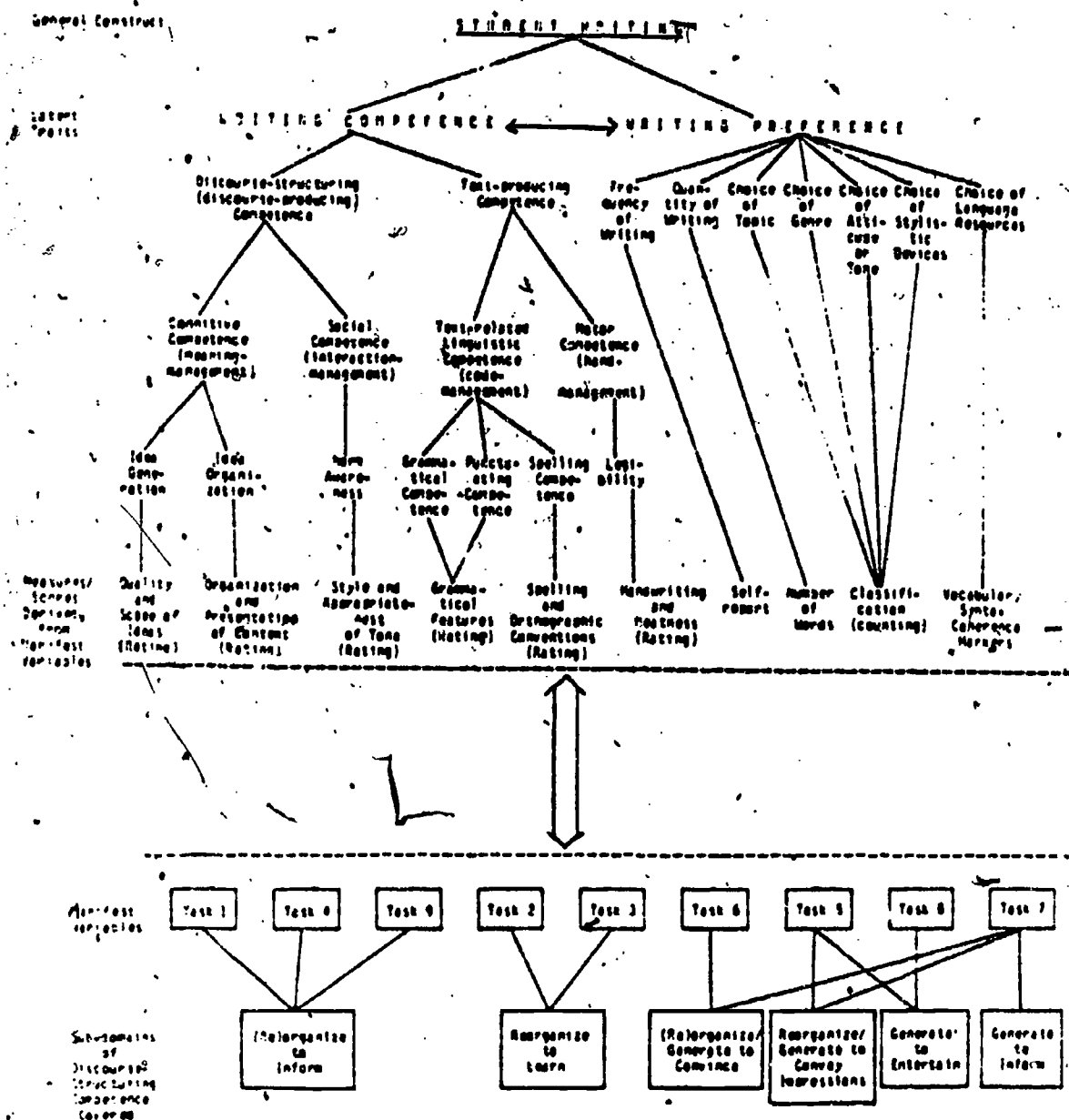


Figure 1. Conceptual structure of writing in relation to scoring system and writing tasks

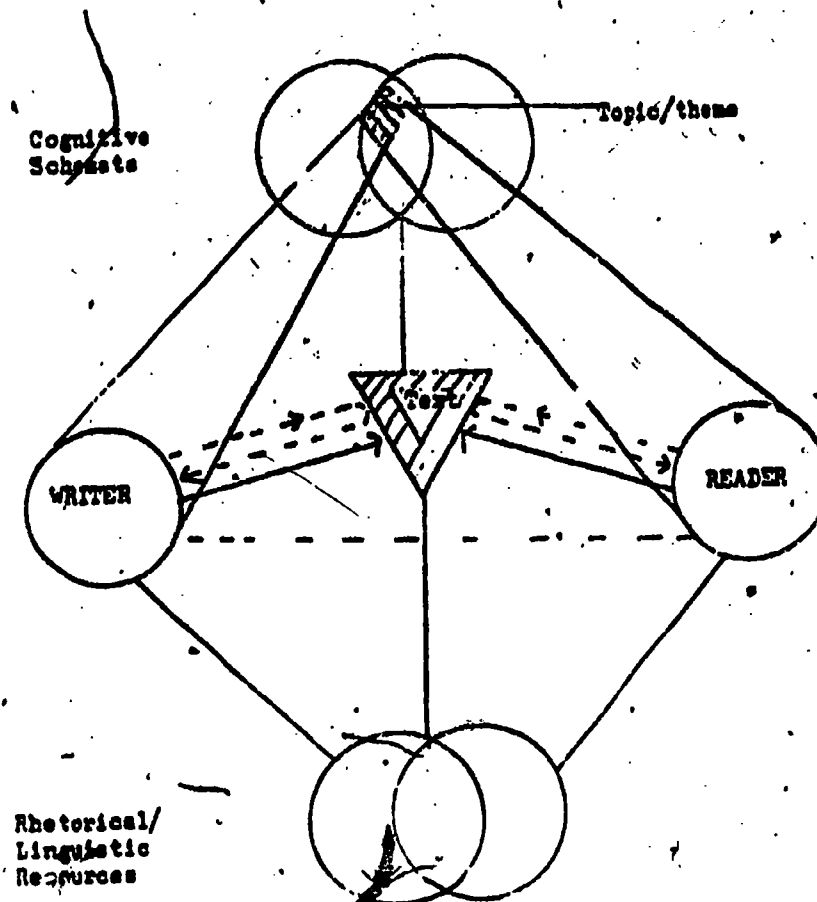


Figure 2. Rhetorical model of writing

TASK 6 - ARGUMENTATIVE/PERSUASIVE COMPOSITION

Population _____ Student _____ Rater _____

WITH RESPECT TO WRITER'S CHOICE
OF AIM AND READERSHIP

Inadequate

Excellent

1 2 3 4 5

Overall Impression

A. Quality and Scope of Content

1. Significance of what is said

2. Argumentation/Exposition

B. Organization and Presentation of Content

3. Organization of the whole text

4. Organization of sub-units

C. Style and Tone

5. Choice of consistency of tone

6. Choice of words and phrases, sentence
structures, and larger units of
discourse

D. Grammatical features

E. Spelling and orthographic conventions

F. Handwriting and neatness

G. Response of Rater

Low Medium High

8. My interest in the composition is

9. My sense of being persuaded by
the composition
